A SURVEY ON BOUNDARY DELINEATION METHODS

Mathias J.P.M. Lemmens
Faculty of Geodetic Engineering
Delft University of Technology
The Netherlands
lemmens@geo.tudelft.nl

Commision III, Working Group 3

KEY WORDS: Photogrammetry, Remote Sensing, Feature, Edge, Extraction, Status, Theory

ABSTRACT

The importance of boundary delineation is indicated by the large amount of literature devoted to the topic. Although subject of intensive research the last three decades the problem is still poorly understood and largely unsolved. Main reasons for failing are that the image models underlying the design of these schemes form a poor description of the actual data set, and that the relationship between data and required information can be modeled only very weakly. The aim of the present paper is to structure the massive volume of edge detection approaches and to arrive at insight into their major merits and shortcomings.

1 Introduction

The role of delineation of boundaries is crucial for a broad range of geo information related activities, such as semi-automatic mapping, GIS-updating, stereo-matching, and object-based multispectral classification. Anyone who has been involved in the extraction of objects from unrestricted scenes, such as recorded by aerial and space imagery, will have encountered difficulties with obtaining reliable object outlinings. Indeed, one of the key problems that makes realization of the above tasks so hard is the outlining of boundaries. Although several attempts have been undertaken to put edge detection in a more rigorous mathematical framework, including: Brooks (1978), Marr & Hildreth (1980), Haralick & Watson (1981), Hildreth (1983), Canny (1986), Nalwa & Binford (1986), and Torre & Poggio (1986), a coherent theory could not be developed. No general algorithms which can be applied successfully on all types of images, have emerged. The relative merits and characteristics of the many individual methods when applied to unrestricted real-world scenes are not at all clear. Numerous legends circulate about the relative merits of different operators (Fleck, 1992). Therefore, the choice of a particular edge detection scheme seems to be more based on the appreciation and preoccupation of the user than on the real capabilities of the scheme. Our aim is to structure the existing methods and to examine their merits, based on our extensive experience on the subject (Lemmens, 1996). Existing surveys can be subdivided into those solely devoted to edge detection, including: (Davis, 1975; Levialdi, 1981; Peli & Malah, 1982) and the ones which discuss segmentation more generally, including: (Fu & Mui, 1981; Haralick & Shapiro, 1985; Pal & Pal, 1993). Furthermore regular textbooks (e.g. Rosenfeld & Kak, 1982; Pratt, 1991; Ballard & Brown, 1982; Davies, 1990) present introductions. Segmentation schemes may be divided into three main categories (Fu & Mui, 1981; Sonka et al., 1993): (1) characteristic feature thresholding or clustering, (2) region extraction, and (3) edge finding. We focus here on edge finding, and more specifically, on local edge detection schemes.

2 Edge Finding Process

Basicly edge finding schemes consist of (1) edge detection, and (2) edge localization. The edge detection part, which is the hard problem and is therefore considered here solely, consists of four steps: (1) smoothing, (2) local edge detection, (3) thresholding and thinning, and (4) edge linking.

In general, local edge detection is based on some form of differentiation of the local grey value function. Since differentiation is a mildly ill-posed problem (Torre & Poggio, 1986) smoothing is often applied beforehand for regularization purposes. Nevertheless, smoothing should be avoided, when possible, since linear smoothing tends (1) to blur the weak edges, (2) to reduce the localization accuracy, and (3) to merge closely spaced edges, while non-linear smoothing filters, such as the Kuwahara and the median filter, tend to dislocate edges and to remove thin, line-shaped objects such as roads. Furthermore, smoothing introduces correlation among the observations which may deteriorate the performance of subsequent processing steps.

Thresholding is a decision process in which the label edge or non-edge is assigned to each pixel, based on the response of the local edge detector. Usually the response is tested against one or more prespecified thresholds. These thresholds may be determined on an heuristical basis or by a quantification of image disturbances such as noise.

Due to the spatial extent of local edge operators, the initial edge map is in general not one pixel thick. Thinning is necessary to obtain one pixel thick ourlinings. One of the possibilities is to use, after thresholding, a skeletonizing algorithm to erode the thick edges. To obtain higher localization precision one may use, before thresholding, non-maximum suppression to exclude a pixel as edge if its edge response is lower than those of the neighbouring pixels located perpendicular to its gradient direction. The disadvantage is that junction pixels may be deleted too. Lacroix (1988) proposes a remedy by allowing edge pixels to form relative maxima, i.e. real edge pixels are permitted to have pixels with higher responses in their vicinity as long as there are sufficient pixels in the neighbourhood with lower responses.

Finally, the edge pixels are linked to form a boundary of connected pixels, that may be generalized and vectorized in a postprocessing stage for storage in, for example, a GIS. To obtain more reliable results one may examine the operator responses in a neighbourhood of connected pixels, using context information.

3 Boundaries

What constitutes a boundary? Since the human visual system is so highly sophisticated, this seems to be a trivial question. However, a large problem in boundary detection is obtaining a suitable definition of a boundary which is generally applicable. Definitions can be given in two domains: (1) image domain, and (2) task domain. Within the image domain the most important feature that may represent a boundary is the edge:

An edge is an intensity discontinuity in the image.

In course of time an abundance of algorithms have been developed, designed to trace local intensity discontinuities in the image. We treat some of them in sections 4 and 5.

The objects one wants to extract, depend on the task domain at hand. For example, for a large scale base map of an urban area one wants other objects than for a national road data base. This observation results in the important understanding that a perfect outlining of relevant objects in aerial and space imagery can not be established by bottom-up approaches alone. We define a boundary as:

A boundary is a (closed) outlining of an object which is relevant for the task at hand.

A common preassumption is that abrupt intensity changes in the image correspond to meaningful object boundaries in the scene. This may be valid for highly restricted scenes, such as present in industrial environments, but for unrestricted scenes, a boundary *may* be visible in the image as abrupt intensity changes, but this is neither a necessary nor a sufficient condition. Not all intensity changes correspond with relevant object outlinings. They may be, for example, due to noise, texture and shadows. Furthermore, boundaries may not show up as intensity changes, due to low contrast or occlusion. To resolve this problem a priori information about the objects relevant to the task domain is necessary. This information can concern radiometric and geometric properties.

The exploration of radiometric properties is carried out regularly by the remote sensing community for multispectral aerial and space imagery by statistical pattern recognition techniques using as features (functions of) the grey values of the different spectral bands. However, pixel-based classification is highly prone to error, resulting in the wish of developing object-based classification techniques, which require, at turn, reliable delineation of objects.

Geometric properties may concern generic or specific aspects. Generic geometric aspects concern descriptions of general appearances of boundaries, such as: (1) smooth continuation, i.e. nearby edge pixels will point approximately in the same direction, (2) thinness, e.g. edges should be one pixel thick, which criterion is often necessary for further processing, and (3) connectivity, e.g. boundaries will be closed. A classical mechanism to incorporate generic geometric information is by probabilistic relaxation (e.g. Schachter et al., 1977; Peleg, 1980; Prager, 1980; Hancock & Kittler, 1990; Duncan & Birkhölzer, 1992). For edge detection purposes on images of unrestricted scenes this mechanism shows several severe drawbacks.

Specific geometric information concerns more detailed descriptions about the shape and possibly size of the boundaries, for example, the objects of interest have circular or rectangular boundaries. The exploration of specific geometric information has been subject of extensive research. For example the Hough transform (for a good survey with many references, see Illingworth & Kittler (1988)), boundary delineation by dynamic programming (e.g. Montanari, 1971; Martelli, 1976; Elliott & Srinivasan, 1981; Gerbrands, 1988; Lemmens et al., 1990), and perceptual grouping (e.g. Witkin & Tenenbaum, 1986; Khan & Gillies, 1992; Lu & Aggerwal, 1992; Mohan & Nevatia, 1992; Price & Huertas, 1992; Lin et al, 1995) are well-known approaches. Lemmens et al. (1990) suggested to use dynamic programming for road detection. Within the photogrammetric community De Gunst et al. (1991), and Grün & Li (1994) explored the approach for this purpose.

We focus in the sequel on local edge detection schemes, because they form the fundamental stage in any boundary delineation task. We loosely divide them into two broad classes: (1) monadic schemes and (2) plural schemes. Monadic schemes base the decision whether an edge is present or not, directly on the responses of the operators, without further examination of additional responses, while plural schemes carry out such an evaluation in some form.

4 Monadic Local Edge Detection

A local edge detector is an operator of small spatial extent that traces changes in the image function to classify each pixel as edge or non-edge according to some decision rule. No a priori information about the scene structure or contextual information is employed. Monadic methods can be generally classified into one of four categories: (1) Differentiation, (2) Surface fitting, (3) Template matching, and (4) Curvature determination. Other approaches have been proposed, such as those based on morphological operators (cf. Lee et al. 1987). We do not treat them here.

4.1 Discrete Differentiation

Abrupt intensity changes are traceable by computing the partial derivatives in two orthogonal directions: $g_x = \partial g/\partial x$; $g_y = \partial g/\partial y$), usually along the grid lines. The smallest step possible on a sampled space is the sample interval Δx. By definition the pixel size is unity: $\Delta x = 1$. Hence the discrete derivative in x-direction of the 2-D discrete function g becomes:

$$g_x = \lim_{\Delta x \to 1}[g(j+\Delta x) - g(j)]/\Delta x = g(j+1) - g(j) \quad (1)$$

Accordingly one can define g_y, the first derivative in y-direction. (For the Roberts operator g_x and g_y become: $g_x = g(i,j+1) - g(i+1,j)$; $g_y = g(i,j) - g(i+1,j+1)$.) The edges located by the gradient components according Eq.(1) are situated at the grid lines yielding interpixel (crack) edges. The computation can be done by convolving g with the masks $[-1\ 1]$ and $[-1\ 1]^T$, where T denotes the transpose. These above masks are not symmetrically positioned. To evade this we may choose the central derivative:

$$g_x = \lim_{\Delta x \to 2}[g(j + \tfrac{1}{2}\Delta x) - g(j - \tfrac{1}{2}\Delta x]/\Delta x$$

which results in: $g_x = \frac{1}{2}[g(j+1) - g(j-1)]$. Accordingly g_y is defined. The computation of g_x and g_y can now be done by convolving g with the masks $[-1\ 0\ 1]$ and $[-1\ 0\ 1]^T$, respectively. The orientation of the edge is defined by: $\theta = \arctan g_y/g_x + 1/2\pi$, and the edge strength (magnitude) by: $M_1 = \sqrt{g_x^2 + g_y^2}$. Alternative definitions of edge strength that bypass squaring and square rooting reducing computation time, are the Manhattan distance: $M_2 = |g_x| + |g_y|$, and the chess-board distance: $M_3 = \max(|g_x|, |g_y|)$. The magnitude and edge direction show a directional bias; they depend on the orientation of the edge with respect to the grid. For example the magnitude error of the Sobel operator may reach 7.93 % and the angle error 2.90 degrees for ideal step edges, while the magnitude error of the Prewitt operator may even reach 12.87 % and the angle error 7.43 degrees (Lyvers & Mitchell, 1988). To replace M_1 by M_2 or M_3 is disadvantageous since the orientation dependent bias is more severe for the last two measures. M_3 corresponds to template matching with just two hypothesized directions (section 4.3). The gradient components $|g_x|$ and $|g_y|$ will occupy only a limited grey value range -typically in the range of the grey values when properly normalized- which is usually in the range 0-255. Therefore, when using the Euclidean distance M_1, one may save computation time by precomputing all edge strengths that may occur and to store them in a look-up table. Although the Prewitt (1970) operator is derived from a surface fitting approach (section 4.2), its 3 × 3 version can also be understood as combining discrete differentiation with an unweighted smoothing perpendicular to the direction over which the gradient is computed, i.e. horizontal gradient component: $[-1\ 0\ 1] * [1\ h\ 1]^T$, and vertical component: $[-1\ 0\ 1]^T * [1\ h\ 1]$, with $h = 1$ and $*$ the convolution sign. For the Sobel-operator $h = 2$ and for the Isotropic operator $h = \sqrt{2}$.

The Sobel operator has been shown to be superior to other small support operators, like the 3 × 3 Prewitt operator. The background of the Sobel operator is that the grey values that lie closest to the central pixel become a higher weight than the grey values which lie farther away, yielding good smoothing properties.

Since also other features than edges show up as abrupt intensity changes, e.g. noise and texture, differentiation approaches actually detect non-edges. When the response exceeds the prespecified threshold, it is assumed that the detected grey value variability is due to the presence of an edge. The early edge detection schemes, such as the Sobel operator, were developed on a more or less intuitive base, without much mathematical foundation. When one wants to detect a feature in a signal in a reliable way, two phenomena should be at least employed: (1) a model of the appearance of the feature in the signal, and (2) knowledge about corruption of the signal with disturbances such as noise. Commonly, edges are modeled as ideal step functions, i.e. two grey values are assumed to be present in the local neighbourhood. The disturbances are usually modeled as additive zero-mean Gaussian distribtued noise, uncorrelated with the signal. The remaining part of this section treats methods based on such image models.

4.2 Surface Fitting

In an attempt to make edge detectors more immune to non-edge features, much research efforts has been devoted to model explicitly how edges may look like in the local image function. One of the results are surface fitting approaches. The local image function is modeled as a set of basis functions, that express the theoretical edge. The surface is fitted to the local image function according some optimalization norm, usually the L^2-norm.

Prewitt (1970) -she was the first to suggest the surface fitting idea- approximates the local image function by a second order polynomial: $g(x, y) = a_0 + a_1x + a_2y + a_3xy + a_4x^2 + a_5y^2$ resulting, after a least squares fit, in a 3 × 3 neighbourhood in the Prewitt masks discussed in the previous section.

The Hueckel (1973) method involves finding the parameters of the best fitting ideal step edge, assuming two grey values in a circular neighbourhood of 32 to 137 pixels. The fitting is done by the L^2-norm using a set of two-dimensional orthonormal basis functions by a Fourier series in polar coordinates. The expansion is truncated to eight terms for computational and smoothing reasons.

O'Gorman (1976) modifies the approach of Hueckel, by using Walsh functions, instead of sin/cos basis functions. The rationale governing this choice is that the discrete image space bears a simple relationship to Walsh functions.

Ghosal & Mehrotra (1993) model an ideal step edge, assuming the presence of two grey values, by a set of orthogonal complex Zernike moments.

Haralick (1984) extents the polynomial approach of Prewitt (1970), by suggested a zero-crossing edge detector based on the facet model, introduced in Haralick (1980), using directed second order derivatives.

4.3 Template Matching

Another approach to explicitly model the appearance of edges is by using templates. Template matching for edge detection purposes is the process of moving a two-dimensional template, representing a prototype edge, over the entire image. At every pixel the local image function over a patch, which has the same extent as the template, is compared with the template. The elements of the templates can be chosen freely, as long as they reflect the underlying edge model. It is convenient to introduce normalized templates, in particular: semi-normalized, normalized, and fully-normalized templates. Let $h_i, i = 1, \dots, n_\kappa$ be the elements of a template $\mathbf{h}$.

A template is *semi-normalized* if the mean of the elements is zero: $\sum_{i=1}^{n_\kappa} h_i = 0$. An example is shown in Figure 1a.

A template is *normalized* if the mean of the elements is zero and its variance is $m/(n_\kappa - 1)$, where m is the number of non-zero elements within the template: $\sum_{i=1}^{n_\kappa} h_i = 0$; $\sum_{i=1}^{n_\kappa} h_i^2 = m$. An example is shown in Figure 1b. This definition yields two consequences: (1) the template elements can only take the values $-1, 0$ and 1, and (2) the number of elements that have value -1 equals the number of elements with value 1. So, the number of same-signed values is $\frac{1}{2}m$.

A template is *fully normalized* if the mean of the elements is 0 and its variance is $1/(n_\kappa - 1)$: $\sum_{i=1}^{n_\kappa} h_i = 0$; $\sum_{i=1}^{n_\kappa} h_i^2 = 1$. An example is shown in Figure 1c. To examine the presence of an edge at each pixel, multiple templates are needed. Each template is associated with a hypothesized edge orientation. Let there be K orientations and consequently K templates:

$$\underset{a}{\begin{vmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{vmatrix}} \quad \underset{b}{\begin{vmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{vmatrix}} \quad \underset{c}{\frac{1}{\sqrt{6}} \begin{vmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{vmatrix}}$$

Figure 1: *Examples of normalized templates: (a) semi-normalized, (b) normalized, (c) fully normalized.*

$\mathbf{h}_k, k = 1, \ldots, K$, then we denote the response of template $\mathbf{h}_k$ to an image patch g containing the grey values g_i by: R_{gh_k}..

Remarks

(1) The widely used edge and line detection templates encountered in literature fulfil the above defined normalization conditions (see, e.g., Nevatia, 1982; Rosenfeld & Kak, 1982; Davies, 1990; Pratt, 1991). For example, the eight masks of the three level Robinson (1977) operator, are obtained by permuting the mask coefficients of Figure 1b cyclically. In the same way, cyclical permutation of the mask coefficients of Figure 1a yields the eight masks of the Kirsch (1971) operator.

(2) Each hypothesized edge direction requires a template. The maximum response of all directional templates at a pixel defines the edge strength at that pixel. The template producing the largest response defines the edge direction.

(3) Because of the condition: $\sum_{i=1}^{n} h_i = 0$, the three types of templates compute derivatives, $\mathbf{h}$ can contain both first and higher order derivatives.

(4) The usual mask size is 3×3 pixels. Larger template masks are less sensitive to noise and provide a denser division of the edge direction compass-card. However, they require more computational effort. Furthermore, when object density is high, the response will be often a merged version of two or more boundaries.

Depending on the underlying image and noise model, test statistics on the template responses are (Lemmens, 1996):

$$\max_{k=1}^{K} |R_{gh_k}| \geq z_\alpha \sigma_n \sqrt{m}$$

$$\max_{k=1}^{K} |R_{gh_k}| \geq t_{\alpha,\nu} \hat{\sigma}_g \sqrt{\frac{m(n-1)}{t^2_{\alpha,\nu} + n - 2}}$$

$$\max_{k=1}^{K} |R_{gh_k}| \geq t_{\alpha,\nu} \sqrt{m} \sqrt{\frac{\hat{\sigma}_\kappa^2 + \hat{\sigma}_\ell^2}{2}}$$

where z_α is the critical value of the z-score, $t_{\alpha,\nu}$ the critical value of the Student's t-score, with α level of significance and ν the degrees of freedom. σ_n^2 is the variance of the image noise, $\hat{\sigma}_g^2$ is the variance of the grey values covered by the template, and $\hat{\sigma}_\kappa^2$ and $\hat{\sigma}_\ell^2$ are the variances of the grey values at each side of the hypothesized boundary. It is assumed that the templates $\mathbf{h}_k, k = 1, \ldots, K$ are normalized. To obtain tests for semi-normalized templates replace the variable m by $\sum_{i=1}^{n_\kappa} h_i^2$; for fully normalized templates this variable should be replaced by 1.

4.4 Curvature Determination

The image function may be looked at as a two-dimensional curved surface in 3-D space. The structure present in a landscape can be categorized into 8 principle surface types (see e.g. Besl & Jain, 1988): (1) plane, (2) peak, (3) pit, (4) ridge, (5) valley, (6) saddle ridge, (7) saddle valley, and (8) minimal. These eight surfaces are uniquely determined by the sign and value of the two principle curvatures.

It can be shown that the principle curvatures κ_1 and κ_2 of $g(x, y)$ can be achieved by solving the quadratic form: $\kappa^2 - \kappa(g_{xx} + g_{yy}) + g_{xx}g_{yy} - g_{xy}^2 = 0$ leading to:

$$\kappa_{1,2} = \frac{g_{xx} + g_{yy}}{2} \pm \sqrt{\left(\frac{g_{xx} - g_{yy}}{2}\right)^2 + g_{xy}^2}$$

According to the definition of discrete differentiation, Eq.(1), using the masks $[-1\ 1]$ and $[-1\ 1]^T$, we obtain:

$$\begin{aligned} g_{xx} &= g(x+1,y) - 2g(x,y) + g(x-1,y) \\ g_{yy} &= g(x,y+1) - 2g(x,y) + g(x,y-1) \\ g_{xy} &= g(x+1,y+1) - g(x+1,y) - g(x,y+1) + g(x,y) \end{aligned}$$

We may also fit through the local image function a second order polynomial (see section 4.2) yielding: $g_{xx} = 2a_4$; $g_{yy} = 2a_5$; $g_{xy} = a_3$, and: $\kappa_1 = a_4 + a_5 + D$; $\kappa_2 = a_4 + a_5 - D$, with $D = \sqrt{(a_4 - a_5)^2 + a_3^2}$. The parameters a_0, a_1, a_2, a_3, a_4 and a_5 can be computed from a least-squares adjustment.

The curvature approach is e.g. used by Dreschler & Nagel (1982) for matching of image time sequences. The method is sensitive to noise and texture due to the need to compute second order derivatives.

4.5 Final Remarks

Although the last three approaches (sections 4.2-4.4) model explicitly the feature to be traced, they are essentially based on differentiation of the local image function. Consequently, the desired immunity to non-edge features is not at all warranted. This yields low performance on images of non-restricted scenes, where many other types of features than edges may be present and where the image function is much more complex than the ideal step edge/Gaussian noise model that underlies the design of the majority of the schemes. Furthermore, the derivation of many operators is done in the continuous domain. Next, the filter is sampled, truncated, and usually implemented with a small local support, often as 3×3 windows. As a consequence, the curious situation may occur that operators that are derived along entirely different theoretical lines, may result in the same convolution filters.

5 Plural Local Edge Detection

The plural methods we consider here are: (1) Marr-Hildreth operator, (2) Canny operator, (3) Förstner operator, (4) Edgeness operator, (5) Cascade of local edge detectors, and (6) Orientation coherence operator.

5.1 Marr-Hildreth Operator

The Marr-Hildreth (1980) operator is not primarily based on any underlying image model but on a theory of the human visual system, based on neurophysiological studies. We consider here only the engineering aspects. The image is first convolved with a Gaussian filter of which the blurring effect is controlled by the scale parameter σ. Next the edges are detected as the zero-crossings of the rotation-invariant Laplacian ($\nabla^2 g = g_{xx} + g_{yy}$). The conjunction of the Gaussian with the Laplacian is called Laplacian of Gaussian (LOG).

The basic notion is that edges appear at a wide variety of scales. Therefore, edges should be detected at several amounts of blur, controlled by σ of the Gaussian. The edges detected at different scales are next combined to form the "primal sketch". This notion has resulted in the more general signal analysis technique of scale space filtering, introduced in the early 1980's by Witkin (1983), and further developed in (Babaud et al. 1986; Bergholm, 1987; Perona & Malik, 1990; Lindeberg, 1990; Zuerndorfer & Wakefield, 1990; Liu et al., 1991; Lu & Jain, 1992). The essential idea is to embed the original image in a family of derived images, the scale-space $g(x, y; \sigma)$ obtained by convolving the original image $g(x, y; 0)$

with a Gaussian kernel $G(x,y;\sigma)$ of standard deviation σ, the scale-space parameter, which describes the current level of scale resolution: $g(x,y;\sigma) = g(x,y,0) * G(x,y;\sigma)$. The larger the value of σ the coarser the resulting resolution and the more the image is blurred. The choice of the Gaussian is motivated by the fact that it is the only kernel in a broad class of functions which satisfies adequate scale-space conditions (Babaud et al. 1986): (1) causality: increase of σ should not generate spurious details and (2) homogeneity and isotropy: the blurring is shift invariant and does not depend upon the grey values. The Marr-Hildreth operator suffers from several deficits:

- Theoretically the zero-crossing contours are closed. However, due to noise, texture and quantization effects, breaks may occur since the magnitude of the pixel differences on the two-sides of a zero-crossing do not exceed an acceptance threshold.
- T-junctions or trihedral vertices are incorrectly detected. Instead of 3 meeting zero-crossing lines 2 disconnected lines are detected, i.e. a spurious line is detected. In general, at positions where the edges are highly curves the zero crossings are located improperly. The larger the width of the Gaussian, the larger this effect.
- Gaussian smoothing causes merging of closely spaced edges, resulting in the detection of phantom edges. For example, a set of parallel edges may be joined to one edge after convolution with the Gaussian.
- A good method to combine the results at different scales is lacking.
- Since the Laplacian is a second derivative operator, the operator is sensitive to noise. (A nonlinear Laplacian for use on noisy images has been developed by van Vliet et al. (1989)).

5.2 Canny Edge Detector

Canny (1986) formulates edge detection as an optimization problem and derives optimal filters for the detection of step edges in the presence of Gaussian noise. The product of SNR and the localization measure in one-dimension is optimized, using as performance criteria: 1) good detection, 2) good localization, and 3) only one response to a single edge should appear. The steps of the scheme are:

1. Filtering of the image to smooth the effects of noise and to produce a multiscale representation of the image data. For computational reasons a suboptimal Gaussian filter is chosen.
2. Differentiation by taking directional first derivatives using templates at an interval of 30^o.
3. Non-maxima suppression by interpolating gradient vectors in a 3×3 neighbourhood.
4. Multithreshold hysteresis linking which draws on information concerning edge-connectivity. Edge-pixels are initially labeled if their response exceeds a high-threshold value. Pixels lying above a weaker response threshold are then admitted provided they belong to edge-segments which are connected to the initially labeled pixels. Finally, unconnected high-response pixels are deleted.

Although the Canny operator may remove genuine high-frequency edge-features such as corners, its performance is good. This is, however, not due to its optimal design for step edges embedded in Gaussian noise, but is based on the use of local context (generic geometric) information in step 4. Based on the Canny approach Petrou & Kittler (1991) developed an optimal detector for ramp edges.

5.3 Förstner Operator

The Förstner (1986, 1993, 1994) operator, which is well-known within the photogrammetric community as interest operator for feature-based matching, is based on examination of a small set of connected gradient components g_x and g_y, enabling to distinguish isotropic structure such as blobs, corners, and texture from non-isotropic structure such as edges and lines (e.g. roads). The operator requires two thresholds: one on the strength of the averaged squared gradients to decide upon the presence of a feature; if a feature is present, the other threshold is to decide whether the feature is isotropic or non-isotropic. Based on this scheme Förstner (1994) developed a framework for low level feature extraction.

5.4 Edgeness Operator

In the template approach (section 4.3), the maximum of the responses is taken as edge measure. The relationships among the directional template responses are not taken into account, although they may give an essential point whether the response is due to noise and texture or to the presence of a real edge pixel. A scheme that is able to carry out such a coherence examination is developed by Cheng (1990). The basic notion of this edgeness operator is as follows. If an edge is present the responses of the templates in subsequent directions, starting from the template oriented along the edge, will be monotonically decreasing and will show symmetry. For noisy pixels this systematics will be absent. If the maximum of the template responses exceeds a threshold and all responses show sufficient systematics, the presence of an edge is accepted. The scheme is especially developed for use on radar images. An extensive analysis carried out in Lemmens (1996) shows that the operator, using 5×5 templates, performs very well on images heavily corrupted by noise and/or texture.

5.5 Cascade of Local Edge Detectors

Usually, one out of the many local edge detectors is applied. However, one may employ several edge detectors in sequence to improve quality and/or to reduce the computational costs. McLean & Jernigan (1988) developed in such away a fast scheme for real time processing of large images. The basic idea is to use in a first stage a time efficient operator that indicates pixels of interest. A simple cross operator over the diagonals (mask: [-1 0 1]) may suffice. This operator may classify non-edge pixels as pixels of interest but the number of edges that are wrongly not detected should be preferable zero, since this type of misclassification can not be corrected in a subsequent stage. Next the pixels of interest are considered more thoroughly by a more sophisticated operator. Tan & Loh (1993) make the above approach still more efficient by using in addition multiple resolutions by establishing an image pyramid, however at the cost of missing closely spaced edges. It will be obvious that cascades of operators may be realized in many ways.

5.6 Edge Orientation Coherence

The basic notion governing edge orientation coherence approaches is that neighbouring pixels positioned on an edge will approximately show equal orientation (cf. Gregson, 1993). Therefore one may examine the gradient directions of the pixels. This can be done along the following line:

1. Compute the mean and the variance of the directions of the gradients in a local neighbourhood (e.g. a 3×3 window) of which the magnitude is above a predefined threshold;
2. Decide whether the orientations of all gradients point sufficiently well into the same direction on basis of the computed variance and an a priori variance measure derived from the edge orientation bias introduced by the detector and a noise estimate, using and F-test.
3. If the computed variance indicates that all orientations are the same, assign to the central pixel the mean of the directions of the gradients.

It is possible to refine the above process, by removing the outliers step by step and by examining whether the remaining orientations point in the same direction and are spatial connected in such a way that they form likely an edge.

6 Discussion

√ The apparently simple problem of locating edges in an image has proved to be very difficult and is still poorly understood. There probably exists virtually no mathematical approach or trick that has been remained untouched to tackle the boundary delineation problem, which is an indication of its intricacy. Optimal methods based on thorough theoretical considerations reveal to produce poor results on aerial and satellite images, due to the fact that the underlying assumptions about the data are often violated. In particular the design of many of the (optimal) edge detection schemes are based on assumptions, which are unrealistic for images of non-restricted scenes, including: (1) the image contains only ideal step edges embedded in zero-mean Gaussian distributed noise, (2) the image may be described as an analytical function, (3) the only intensity changes are locally straight step edges, (4) intensity varies linearly in the direction perpendicular to the edge, (5) edges are broadly spaced, and (6) abrupt intensity changes in the image correspond to meaningful object boundaries in the scene. One of the main reasons for failing is that local edge detectors can not discriminate among the many types of features that may be present in the image. Even in noisy and texture areas, high responses will occur.

√ The above weaknesses of edge detection schemes combined with the fact that the boundary delineation problem is task-domain dependent results in the inevitable conclusion that the exploration of specific geometric object information is indispensable to arrive at reliable boundary outlinings. This conclusion introduces questions like: how to obtain adequate descriptions of specific geometric constraints?, and how to match these constraints with the image function?

√ The main reasons why so many edge detection schemes could emerge, are: (1) the broad variety of mathematical principles and tricks that can be used to base an edge detector on, and (2) existing techniques are often not suited for the particular task the researcher has at hand, forcing to search for other methods resulting in a new approach.

√ It is remarkable that the performance of the many local edge detectors, whether they are based on heuristic grounds or on rigorous mathematical considerations, does not exhibit expressive differences. The choice of the type of preprocessing (smoothing) and the type of postprocessing, in particular context incorporation, reveals to be actually more important for the final result than the choice of a particular local edge detector.

REFERENCES

[1] *Ballard, D.H., Brown, C.M., 1982, Computer Vision*, Prentice-Hall.

[2] *Babaud, J., Witkin, A.P., Baudin, M., Duda, R.O., 1986*, Uniqueness of the Gaussian kernel for scale-space filtering, PAMI, 8 (1), 26-33.

[3] *Bergholm, F., 1987*, Edge focusing, PAMI, 9 (6), 726-741.

[4] *Besl, P.J., Jain, R.C., 1988*, Segmentation through variable-order surface fitting, PAMI, 10 (2), 167-192.

[5] *Brooks, M.J., 1978*, Rationalizing edge detectors, CGIP, 8, 277-285.

[6] *Canny, J., 1986*, A computational approach to edge detection, PAMI, 8 (6), 679-698.

[7] *Cheng, X.S., 1990, Design and implementation of an image understanding system: DADS*, Ph.D. thesis, DUT.

[8] *Davies, E.R., 1990*, Machine vision, AP.

[9] *Davis, L.S., 1975*, A survey of edge detection techniques, CGIP, 4, 248-270.

[10] *De Gunst, M.E., Han, C.S.L.A., Lemmens, M.J.P.M., Van Munster, R.J., 1991*, Automatic extraction of roads from SPOT images, In: OEEPE Official Pub. (27), pp. 131-140.

[11] *Dreschler, L., Nagel, H.-H., 1982*, Volumetric model and 3D trajectory of a moving car derived from monocular TV frame sequences of a street scene, CGIP, 20, 199-228.

[12] *Duncan, J.S., Birkhölzer, T., 1992*, Reinforcement of linear structure using parametrized relaxation labeling, PAMI, 14 (5), 502-515.

[13] *Elliott, H., Srinivasan, L., 1981*, An application of dynamic programming to sequential boundary estimation, CGIP, 17, 291-314.

[14] *Fleck, M.M., 1992*, Some defects in finite-difference edge finders, PAMI, 14 (3), 337-345.

[15] *Förstner, W., 1986*, A Feature based correspondence algorithm for image matching, PRS, 26-III, 1-17.

[16] *Förstner, W., 1993*, Feature extraction in digital photogrammetry, *Photogrammetric Record*, 14 (82), 595-611.

[17] *Förstner, W., 1994*, A framework for low level feature extraction, in: *Lecture Notes in Computer Science, 801*, Springer, 383-394.

[18] *Fu, K.S., Mui, J.K., 1981*, A survey on image segmentation, PR, 13, 3-16.

[19] *Gerbrands, J.J., 1988, Segmentation of noisy images*, Ph.D. thesis, DUT.

[20] *Ghosal, S., Mehrotra, R., 1993* Orthogonal moment operators for subpixel edge detection, PR 26, nr. 2, 295-306.

[21] *Gregson, P.H., 1993*, Using angular dispersion of gradient direction for detecting edge ribbons, PAMI, 15 (7), 682-696.

[22] *Grün, A., Li, H., 1994*, Semi-automatic road extraction by dynamic programming, PRS, 30-III, 324-332.

[23] *Hancock, E.R., Kittler, J., 1990*, Edge-labeling using dictionary-based relaxation, PAMI, 12 (2), 165-181.

[24] *Haralick, R.M., 1980*, Edge and region analysis for digital image data, CGIP, 12, 60-73.

[25] *Haralick, R.M., 1984*, Digital step edges from zero crossing of second directional derivatives, PAMI, 6 (1), 58-68.

[26] *Haralick, R.M., Shapiro, L.G., 1985*, Image segmentation techniques: a survey, CVGIP, 29, 100-132.

[27] *Haralick, R.M., Watson, L., 1981*, A facet model for image data, CGIP, 15, 113-129.

[28] *Hildreth, E.C., 1983*, The detection of intensity changes by computer and biological vision systems, CVGIP, 22, 1-27.

[29] *Hueckel, M.H., 1973*, A local visual operator which recognizes edges and lines, *J. Ass. for Computing Machinery*, 20 (4), 634-647.

[30] *Illingworth, J., Kittler, J., 1988*, A survey of the Hough transform, CVGIP, 44, 87-116.

[31] *Khan, G.N., Gillies, D.F., 1992*, Extracting contours by perceptual grouping, IVC, 10 (2), 77-88.

[32] *Kirsch, R.A., 1971*, Computer determination of the constituent structure of biological images, *Computers and Biomedical Research*, 4 (3), 315-328.

[33] *Lacroix, V., 1988*, A three-module strategy for edge detection, PAMI, 10 (6), 803-810.

[34] *Lee, J.S.J., Haralick, R.M., Shapiro, L.G., 1987*, Morphologic edge detection, *IEEE J. of Robotics and Automation*, 3 (2), 142-156.

[35] *Lemmens, M.J.P.M., 1996*, *Structure-based edge detection*, Ph.D. Thesis, DUT, in preparation.

[36] *Lemmens, M.J.P.M., Meijers, P.J.M., Verwaal, R.G., 1990*, GIS-guided linear feature extraction, PRS, 28, 3/2, comm. III, 395-404.

[37] *Levialdi, S., 1981*, Finding the edge, in: *Digital Image Processing*, Reidel, 383-397.

[38] *Lin, C. Huertas, A., Nevatia, R., 1995*, Detection of buildings from monocular images, in: *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhäuser, 125-134.

[39] *Lindeberg, T., 1990*, Scale-space for discrete signals, PAMI, 12 (3), 234-254.

[40] *Liu, Z-Q., Rangayyan, R.M., Frank, C.B., 1991*, Directional analysis of images in scale space, PAMI, 13 (11), 1185-1192.

[41] *Lu, H.Q., Aggarwal, J.K., 1992*, Applying perceptual organization to the detection of man-made objects in non-urban scenes, PR, 25 (8), 835-853.

[42] *Lu, Y., Jain, R.C., 1992*, Reasoning about edges in scale space, PAMI, 14 (4), 450-468.

[43] *Lyvers, E.P., Mitchell, O.R., 1988*, Precision edge contrast and orientation estimation, PAMI, 10 (6), 927-937.

[44] *Marr, D., Hildreth, E., 1980*, Theory of edge detection, *Proc. R. Soc. Lond.*, B 207, 187-217.

[45] *Martelli, A., 1976*, An application of heuristic search methods to edge and contour detection, *Comm. of the ACM*, 19 (2), 73-83.

[46] *McLean, G.F., Jernigan, M.E., 1988*, Hierarchical edge detection, CVGIP, 44, 350-366.

[47] *Mohan, R., Nevatia, R., 1992*, Perceptual organization for scene segmentation and description, PAMI, 14 (6), 616-635.

[48] *Montanari, U., 1971*, On the optimal detection of curves in noisy pictures, *Comm. of the ACM*, 14 (5), 335-345.

[49] *Nalwa, V.S., Binford, T.O., 1986*, On detecting edges, PAMI, 8 (6), 699-714.

[50] *Nevatia, R., 1982*, *Machine Perception*, Prentice Hall.

[51] *O'Gorman, F., 1978*, Edge detection using Walsh functions, *Artificial Intelligence*, 10, 215-223.

[52] *Pal, N.R., Pal, S.K., 1993*, A review on image segmentation techniques, PR, 26, 1277-1294.

[53] *Peleg, S., 1980*, A new probabilistic relaxation scheme, PAMI, 2 (4), 362-369.

[54] *Peli, T., Malah, D., 1982*, A study of edge detection algorithms, CGIP, 20, 1-21.

[55] *Perona, P., Malik, J., 1990*, Scale-space and edge detection using anisotropic diffusion, PAMI, 12 (7), 629-639.

[56] *Petrou, M., Kittler, J., 1991*, Optimal edge detectors for ramp edges, PAMI, 13 (5), 483-491.

[57] *Prager, J.M., 1980*, Extracting and labeling boundary segments in natural scenes, PAMI, 2 (1), 16-26.

[58] *Pratt, W.K. 1991*, *Digital Image Processing*, Wiley.

[59] *Prewitt, J.M.S., 1970* Object enhancement and extraction, in: *Picture Processing and Psychopictorics*, AP, 75-149.

[60] *Price, K., Huertas, A., 1992*, Using perceptual grouping to detect objects in aerial scenes, PRS, 29-III, 842-855.

[61] *Robinson, G.S., 1977*, Edge detection by compass gradient masks, CGIP, 6, 492-501.

[62] *Rosenfeld, A., Kak, A.C., 1982*, *Digital picture processing*, AP.

[63] *Schachter, B.J., Lev, A., Zucker, S.W., Rosenfeld, A., 1977*, An application of relaxation methods to edge reinforcement, *IEEE Trans. on Systems, Man, and Cybernetics*, 7 (11), 813-816.

[64] *Sonka, M., Hlavac, V., Boyle, R., 1993*, *Image processing, analysis and machine vision*, Chapman & Hall.

[65] *Tan, C.L., Loh, S.K.K., 1993*, Efficient edge detection using hierarchical structures, PR, 26 (1), 127-135.

[66] *Torre, V., Poggio, T.A., 1986*, On edge detection, PAMI, 8 (2), 147-163.

[67] *Van Vliet, L.J., Young, I.T., Beckers, G.L., 1989*, A nonlinear laplace operator as edge detector in noisy images, CVGIP, 45, 167-195.

[68] *Witkin, A.P., 1983*, Scale-space filtering, in: *Proc. 8th Int. Conf. Artificial Intelligence*, Morgan Kaufmann, pp. 1019-1022.

[69] *Witkin, A.D., Tenenbaum, J.M., 1986*, On perceptual organization, in: *From Pixels to Predicates*, Ablex, 149-169.

[70] *Zuerndorfer, B., Wakefield, G.H.*, Extensions of scale-space filtering to machine-sensing systems, PAMI, 12 (9), 868-882.

Used Abbreviations in Reference List	
AP:	Academic Press
CGIP:	Computer Graphics and Image Processing
CV:	Int. Journal of Computer Vision
CVGIP:	Computer Vision, Graphics and Image Processing
DUT:	Delft University of Technology
IVC:	Image and Vision Computing
PAMI:	IEEE Trans. on Pattern Analysis and Machine Intelligence
PR:	Pattern Recognition
PRS:	Int. Arch. of Photogramm. and Remote Sens.